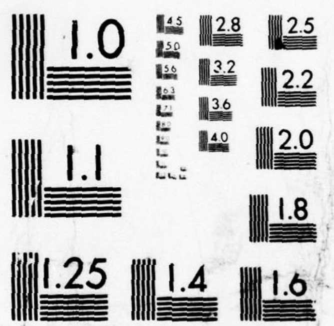
AD-A068 724 HARVARD UNIV CAMBRIDGE MA CENTER ON DECISION AND CON--ETC F/G 12/1
MONITORING COOPERATIVE AGREEMENTS BETWEEN PRINCIPALS AND AGENTS--ETC(U)
FEB 79 R RADNER N00014-77-C-0533

UNCLASSIFIED TR-3 NL

1 OF 1

AD
A068724

END
DATE
FILMED
6 --79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MONITORING COOPERATIVE

AGREEMENTS BETWEEN

PRINCIPALS AND AGENTS

Roy Radner

DDC
RECEIVED
MAY 18 1979
A

Technical Report No. 3

Prepared under Contract No. N00014-77-C-0533
Project No. NR 277-240
for the Office of Naval Research

This document has been approved for public release and sale; its distribution is unlimited.

Reproduction in whole or part is permitted for any purpose of the United States Government.

Harvard University
Littauer #308
Cambridge, Mass. 02138

February, 1979

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

MONITORING COOPERATIVE AGREEMENTS BETWEEN PRINCIPALS AND AGENTS

Roy Radner

1. Introduction[1]

Theories of agency and of the design of incentives in organizations typically portray the members of the organization as players in a noncooperative game. The predictive theory that naturally accompanies this point of view is that of Nash equilibria, including Harsanyi's elaboration of that theory to accommodate situations in which the players have incomplete information about the parameters of the game.

On the other hand, much normative theory of organizations uses the framework of cooperative game theory, with its array of alternative "solution" concepts (value, core, von Neumann-Morgenstern solution, Nash bargaining solution, etc.). Furthermore, empirical observations of organizations reveal widespread cooperative behavior, as well as noncooperative behavior, so that cooperative game theory may have descriptive as well as normative value.

What determines whether members of an organization cooperate or not? Conventional wisdom suggests that cooperation is less likely--or less stable--the more players there are, or the greater the difficulty of communication among the players; cooperation is more likely (stable?) if there are mechanisms whereby the players make binding commitments. Thus theories of industrial

---

[1] This research was supported by the Office of Naval Research, Contract No. N00014-77-C-0533 and by the National Science Foundation (Grant SOC76-14768 to the University of California, Berkeley). A preliminary version of this paper was presented at the CEME-NBER Conference on Decentralization, University of California, San Diego, Feb. 23-25, 1979.

organization typically _assume_ that when the number of firms in an industry is "large" the resulting equilibrium will be of the noncooperative type, whereas when the number of firms is "small" the outcome may be cooperative (collusive).

The theory of repeated games explores in a formal way another piece of conventional wisdom, namely that when members of an organization have long-lasting relationships they can encourage and maintain cooperative behavior (without the device of binding commitments) by signalling intentions to cooperate and by punishing defectors from informal agreements. Indeed, the theory of repeated games provides conditions under which _noncooperative_ equilibria of the entire sequential game can produce _cooperative_ outcomes of the component subgames.

Unfortunately, such results seem to require an infinite number of repetitions of the subgame; they are not valid for a finite number of repetitions, no matter how large that finite number. However, similar results can be obtained for _approximate_ noncooperative equilibria in the finite-repetitions case; such an approximate equilibrium is called an _epsilon_-equilibrium if each player's sequential strategy is within epsilon (in utility) of being the best response to the other players' strategies. Thus, one gets the result that, for any fixed positive epsilon, if the number of repetitions is large enough then there are noncooperative epsilon-equilibria that have cooperative outcomes in each subgame. In a sense, in finite repetitions of a game, the best is the enemy of the good!

In the principal-agent model, the agent observes a (random) environmental variable and then chooses an action; this leads to an outcome that depends on both the action and the environment. The principal observes this outcome (but neither the agent's action nor the environment), and pays the agent according

to a previously announced reward function, which depends on the outcome only.

In equilibria of repeated games that sustain cooperative behavior, each player is "punished" by the others for departures from the informal agreement to cooperate.[2] However, in the principal-agent situation, the principal cannot observe the agent's behavior directly, but only the consequences of his behavior, and those consequences are also influenced by the environment. Therefore, if cooperative agreements are to be sustained as equilibria of the repeated game, the principal must have some statistical method of detecting "cheating" by the agent rapidly enough to deter him from doing so; on the other hand, this method should have a very low probability of triggering false alarms. The main theorem of this paper (Sec. 5) shows that this is possible.[3]

In Sections 2 and 3, I present the principal-agent model in the form of a one-period game, and state a few of its properties. In Section 4 I review the essential concepts in the theory of epsilon-equilibria of finitely repeated games. Section 5 contains the main result on the existence of epsilon-equilibria in T-period repetitions of the principal-agent game, when T is large (but finite). The proof is constructive, and exhibits a family of epsilon-equilibrium strategy pairs. Using this family of strategy pairs one can approach

---

[2] An early important paper on repeated games (supergames) is by Aumann (1959). Characterizations of perfect Nash equilibria in infinite supergames have been provided by Aumann and Shapley (unpublished) and by Rubinstein (1977). For an analysis of altruism in the context of infinite supergames see Kurz (1978). Examples of epsilon-equilibria of finite supergames have been studied by Radner (1979a, 1979b).

[3] The main theorem uses, among other facts of probability theory, the law of the iterated logarithm, and is related to sequential tests of hypotheses that have power one (see Robbins and Siegmund, 1974, and the references given there). Since the research for the present paper was completed, I had the opportunity to see an unpublished paper by A. Rubinstein (1978), in which he uses the law of the iterated logarithm to demonstrate the existence of Nash equilibria with close to Pareto optimal average expected utility in an example of an infinite supergame.

arbitrarily close, in terms of average expected utility per period, to any one-period cooperative arrangement that dominates a one-period Nash equilibrium. Section 6 indicates some extensions of the theory.

2. A Model of a Sequential Principal-Agent Relationship

Consider a principal-agent relationship that lasts T periods. In period t, the agent's action is A_t, a number between 0 and M_a (a positive parameter). The outcome of the agent's action is

$$C_t = \gamma(A_t, Z_t) ,$$

where Z_t is an exogenous random variable (the "state of nature" in period t). We may interpret the variable A_t as a measure of the agent's effort. The principal observes the outcome of the agent's action, and pays the agent W_t. The resulting one-period utility to the agent is $U(W_t, A_t)$, where the function U is strictly concave, increasing in W, and decreasing in A. The one-period utility to the principal is assumed to be a linear function of the outcome and the payment to the agent, increasing in the former and decreasing in the latter. By a suitable choice of units one can express the principal's utility as $C_t - W_t$. The agent can observe the state of nature, A_t, before taking action, but the principal can observe only the resulting outcome, C_t.

Assume that the functions U and γ are continuously differentiable, that for every Z the function $\gamma(\cdot,Z)$ is concave and increasing in its first argument (the agent's action), and that the partial derivative of γ with respect to the agent's action is bounded away from 0, uniformly in Z, say $\geq M' > 0$.

Notice that I have assumed that the agent is risk-averse, whereas the principal is risk-neutral. The main theorem (Section 5) can easily be extended to the case in which the principal is risk-averse; see Section 6.

3. The One-Period Game

In this section I review the usual formulation of the principal-agent relationship as a one-period noncooperative game.[4] I therefore omit the subscript t on all the variables. The principal's (pure) strategy is a reward function ω that determines the payment to the agent as a function of the outcome of the agent's action :

$$W = \omega(C) \quad .$$

Given the reward function ω, the agent chooses a decision function α that determines his action as a function of the state of nature:

$$A = \alpha(Z) \quad .$$

The expected utility to the agent is

$$\mathcal{E}U\{\omega(\gamma[\alpha(Z),Z]), \alpha(Z)\}$$

and the expected utility to the principal is

$$\mathcal{E}\gamma[\alpha(Z), Z] - \mathcal{E}\omega(\gamma[\alpha(Z), Z]) \quad .$$

This is in fact a two-move game with perfect information, in which the principal moves first, choosing the reward function, and the agent moves second, choosing the decision function. The noncooperative solution to the game is taken to be a Nash equilibrium.

Recall that a pair (ω, α) of functions is Pareto-optimal if there is no other pair that yields each player at least as high an expected utility, and yields at least one of the players strictly more.

Note that the decision function α is a <u>move</u>, not a strategy. The agent's strategy is a mapping from reward functions ω to decision functions α, since the agent learns the reward function before choosing the decision function.

---

[4] For material on the principal-agent problem, see Shavell (1978) and the references cited there. For a more general organizational setting of the problem, see Groves (1973).

For the purposes of this paper, the important characteristics of Nash equilibria and Pareto-optima of the one-period game are summarized as follows.

Proposition

(1) In a Nash equilibrium, the reward function must be strictly increasing on the set of realizable outcomes, i.e., on the range of $\gamma[\alpha(\cdot),\cdot]$.

(2) In a Pareto optimum, the reward function must be constant on the set of realizable outcomes; hence

(3) A Nash equilibrium cannot be Pareto-optimal.

Note that a consequence of property (2) is that, if (ω, α) is Pareto-optimal, then the agent's best response to the (constant) reward function ω is to always set his action equal to 0. In other words, with a reward that is independent of the outcome of the agent's action, he has an incentive to reduce his effort below the level called for by the decision function α.

4. Epsilon-Equilibria of Repeated Games

Suppose now that the one-period game is repeated T times (T finite); the resulting sequential game will be called the T-period game. Assume that the utility to a player is the average of the T one-period expected utilities. A pure sequential strategy for a player is a sequence of functions, one for each period; the function for period t determines the player's one-period strategy in period t as a function of all of the information available to the player up to that period. A Nash equilibrium of the sequential (T-period) game is a pair of sequential strategies such that each player's sequential strategy is a best response to the other player's sequential strategy. Equilibrium pairs of strategies will typically involve threats of "punishment" by one player if the other player departs from some prescribed sequential strategy.

The concept of perfect equilibrium of the T-period game has been introduced by Selten (1975) to rule out equilibria in which the players use threats that are not "credible." For any date and any history of observations up to that date, a player's sequential strategy determines a sequential strategy for the remaining T-t+1 periods of play, which we may call the continuation of the original sequential strategy, given the period and the history of observations prior to that period. A pair of sequential strategies is a perfect Nash equilibrium of the T-period game if, for every period t and every history of prior observations, the respective continuations form a Nash equilibrium of the remaining (T-t+1)-period game. Note that in the definition of a perfect Nash equilibrium one must test, for each period t, whether the pair of continuations is a Nash equilibrium for all possible pairs of prior histories of observations, not just those that would be produced by the original strategy pair.

In games in which each player can observe the other player's one-period strategies (and not just the consequences of action), one can show that in every perfect Nash equilibrium of the T-period game, the strategy pair used in every period is a Nash equilibrium of the one-period game. On the other hand, one can show that, if T is infinite, there are perfect equilibria of the sequential game that result in the use of "cooperative" pairs of strategies in each one-period game, and in particular in the use of Pareto-optimal pairs of strategies.[5] This discontinuity at infinity motivates the definition of epsilon-equilibria in the T-period game (T finite). (See Radner, 1979a and 1979b.) For any positive number epsilon, an epsilon-equilibrium is a pair of strategies such that each player's strategy is within epsilon in average expected utility of being a best response to the other player's strategy. The concept of perfect Nash equilibrium can be extended to epsilon-equilibria as follows. A sequential strategy pair is a perfect epsilon-equilibrium if, for every period t and every history of prior observations, the continuation of each player's strategy is within epsilon of being the best response to the corresponding continuation of the other player's strategy. In this definition, the utility of a continuation of a strategy is the average of the player's expected utilities in all T periods. (For an alternative definition, see Section 6.)

For games in which each player can observe the other player's one-period strategies, one can show that, for any positive epsilon, if T is sufficiently large then there are perfect epsilon-equilibria of the T-period game that result in Pareto-optimal strategy pairs in each one-period game. In other

---

[5] R. Aumann and L. Shapley, unpublished.

words, for perfect epsilon-equilibria, infinite-horizon games are approximated well by long finite-horizon games.[6]

Cooperative one-period strategies can be sustained in perfect epsilon-equilibria of the T-period game by "trigger strategies." Let (s_1^*, s_2^*) be a Nash equilibrium of the one-period game, and let (s_1, s_2) be a Pareto-superior pair of one-period strategies. A trigger strategy for player 1 is defined as follows: player 1 plays strategy s_1 as long as player 2 plays strategy s_2; thereafter player 1 plays s_1^*. The best response by player 2 to this trigger strategy is to play s_2 until the last period, and then play a best response to s_2. However, the gain in average per-period utility of doing this, over using the corresponding trigger strategy, will be small if T is large.

The efficacy of such simple trigger strategies in sustaining perfect epsilon-equilibria of the T-period game depends on each player being able to rapidly detect departures from the cooperative strategies. In the principal-agent situation considered in this paper, the principal cannot observe the agent's actions directly, but only the consequences of his actions, and these consequences also depend on a random state of nature. Therefore, if cooperative arrangements are to be sustained as equilibria of the T-period game, the principal must have available some more powerful method of detecting "cheating" by the agent rapidly enough to reduce the agent's incentive to cheat to negligible levels. That such a method exists is shown in the next section.

---

[6] These results are illustrated in (Radner, 1979a, 1979b). A more general treatment of epsilon-equilibria will be presented in a forthcoming paper.

5. Epsilon-Equilibria of the T-Period Principal-Agent Game

Let (ω^*, α^*) be a Nash equilibrium of the one-period principal-agent game, and let $(\hat{w}, \hat{\alpha})$ be a Pareto-superior pair, where $\hat{w}$ is constant. In this section I shall exhibit a class of perfect epsilon-equilibria of the T-period game, using trigger-type strategies.

Defining a trigger strategy for the agent presents no problem; the agent uses the decision function $\hat{\alpha}$ until the first time the principal does not use the constant reward $\hat{w}$, and then optimizes against the announced reward functions from that period on. I shall denote this strategy by σ_A.

It is important to emphasize at this point that in each one-period game the principal's action is an announcement of a reward function, and he is required to use that reward function for that period. The agent then observes the current Z_t, and takes an action, A_t.

Defining a suitable trigger strategy for the principal is more difficult. In each period t, based on the history of outcomes $C_1, \ldots, C_t$, the principal must decide whether to make the payment $\hat{w}$ or to switch to the Nash equilibrium reward function ω^*. If his switching rule is too lax, then the agent may be able to accumulate a large enough extra expected utility by cheating before getting caught so as to make cheating attractive. On the other hand, if the switching rule is too strict (too "trigger happy"!), then there will be a substantial probability that the principal will switch to the Nash equilibrium reward function before the agent ever starts cheating.

For the remainder of this paper, assume that the states of nature, Z_t, are independently and identically distributed, and bounded. Define

$$\hat{C}_t \equiv \gamma[\hat{\alpha}(Z_t), Z_t] ;$$

thus $\hat{C}_t$ is the realized consequence in period t if the agent uses the decision rule $\hat{\alpha}$. The $\hat{C}_t$ are independently and identically distributed, and bounded; let $\hat{c}$ denote the expected value of $\hat{C}_t$.

Whatever the sequential strategies actually used by the players, let C_t denote the realized outcome in period t, and let $S_n = C_1 + \ldots + C_n$. The C_t are bounded, say by M. Let (b_n) be a strictly increasing sequence of positive numbers ($n \geq 1$), and define the random variables $\tilde{N}$ and N by:

$$
\begin{aligned}
&\tilde{N} = \min \{n \geq 1 : S_n - n\hat{c} \leq -b_n\} , \\
(5.1)\qquad & \\
&N = \min \{\tilde{N}, T\} .
\end{aligned}
$$

Consider the following trigger strategy for the principal: pay the agent $\hat{w}$ in each period 1 through N, and thereafter use the reward function ω^*. I shall denote this strategy by $\sigma_P((b_n))$.

Define $\hat{S}_n = \hat{C}_1 + \ldots + \hat{C}_n$. If the agent uses some sequential strategy other than σ_A, then the principal's loss (if any) during periods 1 through N is

$$L_N \equiv \hat{S}_N - S_N .$$

<u>Lemma 1</u>. If the principal uses the trigger strategy $\sigma_P((b_n))$, then a bound on his expected loss[7] during periods 1 through N is given by

$$\mathcal{E} L_N \leq \mathcal{E} b_N + M \leq b_T + M .$$

<u>Proof</u>. $(\hat{S}_n - n\hat{c})$ is a martingale, and N is a bounded stopping time. Hence, by the systems theorem for martingales (see, e.g., Chung, 1974),

[7] Although it is convenient to interpret L_N as the principal's loss, this is not essential to the argument that follows. What is essential is that this is the cumulated difference in outcome in the direction of the <u>agent's gain</u>. This will become clear in Lemma 2.

$$(5.2) \qquad \mathcal{E}(\hat{S}_N - n\hat{c}) = \mathcal{E}\hat{S}_N - \hat{c}\mathcal{E}N = 0 \ .$$

By the definitions of N and L_n,

$$\hat{S}_N - L_N - N\hat{c} \geqq -b_N - M \ ;$$

taking the expected value of both sides of the above yields

$$(5.3) \qquad \mathcal{E}\hat{S}_N - \mathcal{E}L_N - \hat{c}\mathcal{E}N \geqq -\mathcal{E}b_N - M \ .$$

Also, since the sequence (b_n) is increasing, $\mathcal{E}b_N \leqq b_T$. Putting this last together with (5.2) and (5.3) yields the conclusion of the lemma

Lemma 1 establishes a limit to the cumulated expected loss that the principal can suffer up through period N. The next lemma establishes a corresponding limit on the agent's gain. Let A_t be the agent's actual action in period t, and let $\hat{A}_t$ denote what his action would be if he used the decision function $\hat{\alpha}$, i.e., $\hat{A}_t = \hat{\alpha}(Z_t)$. The corresponding difference in the agent's utility is

$$D_t = U(\hat{w}, A_t) - U(\hat{w}, \hat{A}_t) \quad ,$$

if the agent receives the payment $\hat{w}$. The agent's total gain in utility during periods 1 through n is

$$G_n = D_1 + \ldots + D_n \ .$$

Recall that it has been assumed that

$$(5.4) \qquad \gamma_1(A, Z) \geqq M' > 0 \quad ,$$

where γ_1 denotes the partial derivative of γ with respect to its first argument. In addition, Z is bounded, so that the regularity properties of U

imply that, for some number $M'' > 0$,

$$U_2(\hat{w}, A) \geq -M'' \quad , \text{ for all } A. \tag{5.5}$$

Lemma 2. If the principal uses the trigger strategy $\sigma_P((b_n))$, then a bound on the agent's possible expected gain in utility up through period N is given by

$$\mathcal{E}G_N \leq \left(\frac{M''}{M'}\right)(\mathcal{E}b_N + M) \leq \left(\frac{M''}{M'}\right)(b_T + M) \quad .$$

Proof. By the concavity of γ in A,

$$C_t - \hat{C}_t \leq (A_t - \hat{A}_t)\gamma_1(\hat{A}_t, Z_t) \quad ,$$

$$A_t - \hat{A}_t \geq \frac{C_t - \hat{C}_t}{\gamma_1(\hat{A}_t, Z_t)} \quad ,$$

$$\hat{A}_t - A_t \leq \frac{\hat{C}_t - C_t}{\gamma_1(\hat{A}_t, Z_t)} \leq \frac{\hat{C}_t - C_t}{M'} \quad ; \tag{5.6}$$

The last inequality follows from (5.4).

Recall that U_2 is negative. By the concavity of U, and by (5.5) and (5.6)

$$\begin{aligned} D_t &\leq (A_t - \hat{A}_t)U_2(\hat{w}, \hat{A}_t) \\ &\leq (\hat{A}_t - A_t)M'' \\ &\leq \left(\frac{M''}{M'}\right)(\hat{C}_t - C_t) \quad . \end{aligned}$$

Hence

$$G_n \leq \left(\frac{M''}{M'}\right)(\hat{S}_n - S_n) \quad .$$

The conclusion of the lemma now follows immediately from Lemma 1.

Lemma 2 shows how to make the principal's trigger strategy strict enough to keep the agent's incentive to cheat small when T is large. It suffices to use a sequence (b_n) such that (b_n/n) approaches zero as n increases without limit. But how can the principal, with the same trigger strategy, keep small the probability that the agent would be "unjustly" punished if he never cheats? The law of the iterated logarithm provides the answer.

Recall that, by the law of the iterated logarithm (see, e.g., Chung, 1968),

$$(5.7) \qquad \liminf_{n \to \infty} \frac{\hat{S}_n - n\hat{c}}{\sqrt{n \ln \ln n}} = -\sqrt{2 \operatorname{Var} \hat{C}_t} \equiv -\lambda_0 ,$$

where Var $\hat{C}_t$ denotes the variance of $\hat{C}_t$. Define

$$X \equiv \inf_{n>2} \frac{\hat{S}_n - n\hat{c}}{\sqrt{n \ln \ln n}} ;$$

Then $X > -\infty$ almost surely. Hence

$$\lim_{\lambda \to \infty} \text{Prob } \{X > -\lambda\} = 1 .$$

Hence the following has been proved:

Lemma 3. For every $\delta > 0$ there exists a $\lambda > 0$ such that

$$\text{Prob } \{\hat{S}_n > n\hat{c} - \lambda\sqrt{n \ln \ln n} \text{ , for all } n > 2\} \geqq 1 - \delta .$$

Define the sequence (b_n^o) by

$$b_1^o \equiv - \text{ess inf } (\hat{C}_1 - \hat{c})$$

$$(5.8) \qquad b_2^o \equiv - \text{ess inf } (\hat{C}_1 + \hat{C}_2 - 2\hat{c})$$

$$b_n^o \equiv \lambda_o \sqrt{n \ln \ln n} \quad , \; n \geqq 3 \; .$$

Note that (b_n^o/n) approaches zero as n increases without limit.

Define B to be the class of positive sequences (b_n) that satisfy:

(5.9) b_n are strictly increasing, and

$$\lim_{n \to \infty} \frac{b_n}{n} = 0 \; ;$$

(5.10) there exists $\lambda > 1$ such that

$$b_n \geqq \lambda b_n^o \; , \; n \geqq 1 \; .$$

In particular, B contains all the sequences (λb_n^o) with $\lambda > 1$.

Let $\hat{v}$ and $\hat{u}$ denote the expected one-period utilities of the principal and agent, respectively, under the pair $(\hat{w}, \hat{\alpha})$.

<u>Theorem</u>. For any $\varepsilon > 0$ there exists a sequence (b_n) in B and a T_ε such that the pair $[\sigma_P((b_n)), \sigma_A]$ is an ε-equilibrium for all $T \geqq T_\varepsilon$, and yields the principal and agent average expected utilities at least $(\hat{v} - \varepsilon)$ and $(\hat{u} - \varepsilon)$, respectively, for all T.

<u>Proof</u>.

Let v^* and u^* denote the expected one-period utilities of the principal and agent, respectively, under the (Nash equilibrium) pair (ω^*, α^*). Consider a pair $[\sigma_P((b_n)), \sigma_A]$ of sequential trigger strategies, with (b_n) in B. The

corresponding average expected utility to the principal is

$$(5.11) \qquad \left(\frac{1}{T}\right)[(\mathcal{E}N)\hat{v} + (T - \mathcal{E}N)v^*] \ .$$

(Use the martingale systems theorem again.)

Recall that $\hat{v} \geq v^*$. Define

$$\delta \equiv \text{Prob } (N < T) \quad ;$$

Then (5.11) is at least as large as

$$(5.12) \qquad (1-\delta)\hat{v} + \delta v^* \ .$$

This is as large as $\hat{v} - \varepsilon$ if

$$(5.13) \qquad \delta \leq \frac{\varepsilon}{\hat{v} - v^*} \ .$$

If the principal were to switch in any period n to a reward function other than the constant $\hat{w}$, then in that period and thereafter the agent would optimize against the announced reward functions; hence in periods n through T it would be optimal for the principal to use the reward function ω^*. Hence the principal's optimal response to the agent's strategy σ_A is to use the constant reward $\hat{w}$ in <u>all</u> periods. The resulting average expected utility to the principal is $\hat{v}$. Therefore, if (5.13) is satisfied, the strategy $\sigma_P((b_n))$ is within ε of being optimal against σ_A.

If the agent follows strategy σ_A against $\sigma_P((b_n))$, then his average expected utility is

$$(5.14) \qquad \left(\frac{1}{T}\right)[(\mathcal{E}N)\hat{u} + (T - \mathcal{E}N)u^*] \ .$$

Since $\hat{u} \geq u^*$, it follows that (5.14) is not less than

(5.15) $$(1-\delta)\hat{u} + \delta u^* ,$$

which is at least $(\hat{u} - \varepsilon)$ if

(5.16) $$\delta \leq \frac{\varepsilon}{\hat{u} - u^*} .$$

If the agent uses some sequential strategy σ instead of σ_A against the principal's strategy $\sigma_P((b_n))$, then his average expected utility is

(5.17) $$\left(\frac{1}{T}\right)\left[\mathcal{E} \sum_{t=1}^{N} U(\hat{w}, A_t) + (T - \mathcal{E}N)u^*\right] ,$$

where N is the stopping time under the agent's strategy σ. If the agent uses σ_A, his average expected utility is, by (5.15), at least

(5.18) $$\hat{u} + \delta(u^* - \hat{u}) .$$

The <u>increment</u> in average expected utility to the agent from using σ instead of σ_A is therefore not more than the difference between (5.17) and (5.18), which can be written as

(5.19) $$\left(\frac{1}{T}\right)\left[\mathcal{E}G_N + (T - \mathcal{E}N)(u^* - \hat{u})\right] + \delta(\hat{u} - u^*) .$$

Using Lemma 2, and recalling that $\hat{u} \geq u^*$, we see that (5.19) is not greater than ε if, for example,

(5.21) $$\delta \leq \frac{\varepsilon}{2(\hat{u} - u^*)} ,$$

(5.22) $$\left(\frac{M''}{M'}\right)\left(\frac{b_T + M}{T}\right) \leq \frac{\varepsilon}{2}$$

Therefore, the proof of the theorem is completed by taking δ to satisfy both (5.13) and (5.21), and T_ε to satisfy (5.22); the latter is of course possible because (b_T/T) approaches zero as T increases without limit.

6. Extensions

In the model set out in Section 2, it was assumed that the principal's utility is a linear function of the outcome of the agent's action and the payment to the agent. With a small change in the hypothesis, the theorem of Section 5 remains true if the principal's utility is a concave function of these two variables, increasing in the first and decreasing in the second. The required change is that the constant payment $\hat{w}$ be replaced by a possibly nonconstant reward function $\hat{\omega}$. With this more general hypothesis about the principal's utility function, one can no longer guarantee that in Pareto-optimal arrangements in the one-period game the reward function is constant (cf. the Proposition of Section 3).

In another paper (1979b) I have examined an alternative definition of perfect epsilon-equilibrium in which the utility of the continuation of a sequential strategy is the average utility in the remaining periods, rather than the average of the utilities in all T periods. This change makes the definition of perfect epsilon-equilibrium more restrictive, in the sense that, for every positive epsilon, the set of perfect epsilon-equilibria is smaller. In the application referred to, one could show that, with this alternative definition, cooperative behavior would break down as the game approached the horizon T. This conclusion, which is in accord with observation and common sense, can probably be extended to the principal-agent model, but I have not carried out the details.

References

1. R. J. Aumann, "Acceptable Points in General Cooperative N-Person Games, in A. W. Tucker and R. D. Luce, eds., Contributions to the Theory of Games, Vol. IV, Annals Math. Stud. No. 40, Princeton University Press, Princeton, 1959, 287-324.

2. K. L. Chung, A Course in Probability Theory, second edition, Academic Press, New York, 1974.

3. T. Groves, "Incentives in Teams," Econometrica, 41 (1973), 617-631.

4. M. Kurz, "Altruism as an Outcome of Social Interaction," American Economic Review, 68 (1978), 216-222.

5. R. Radner, "Can Bounded Rationality Resolve the Prisoner's Dilemma?" Bell Telephone Laboratories Discussion Paper, March 1979.

6. __________, "Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives," Bell Telephone Laboratories Discussion Paper, March 1979.

7. H. Robbins and D. Siegmund, "The Expected Sample Size of Some Tests of Power One," Annals of Statistics, 2 (1974), 415-436.

8. A. Rubinstein, "Equilibrium in Supergames," Center Res. Math. Econ. Game Theory, Res. Memorandum No. 25, Hebrew University, May 1977.

9. __________, "Offenses that May Have Been Committed by Accident--An Optimal Policy of Retribution," paper presented at the Int. Conference on Applied Game Theory, Vienna, June 13-36, 1978.

10. S. Shavell, "Rick Sharing and Incentives in the Principal and Agent Relationship," Harvard Institute of Economic Research, Discussion Paper No. 619, July 1978 (to appear in the Bell Journal of Economics).

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Center on Decision and Conflict in Complex Organizations	2a. REPORT SECURITY CLASSIFICATION Unclassified
	2b. GROUP

3. REPORT TITLE
6 Monitoring Cooperative Agreements Between Principals and Agents.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report No. 3

5. AUTHOR(S) (First name, middle initial, last name)
10 Roy/Radner

6. REPORT DATE 11 Feb ~~ruary, 19~~79	7a. TOTAL NO. OF PAGES 20	7b. NO. OF REFS 10
8a. CONTRACT OR GRANT NO. 15 N00014-77-C-0533, NSF-SOC76-24768	9a. ORIGINATOR'S REPORT NUMBER(S) 9 Technical Report, No. 3	
b. PROJECT NO. NR-277-240		
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d.		

10. DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its publication is unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

11. SUPPLEMENTARY NOTES 12 23 p. 14 TR-3	12. SPONSORING MILITARY ACTIVITY Logistics and Mathematics Statistics Branch, Department of the Navy, Office of Naval Research, Wash. D.C.

13. ABSTRACT

The situation in which the principal-agent relationship is repeated finitely many times (T) is formulated as a sequential game. For any cooperative arrangement in the one-period game that dominates a one-period Nash equilibrium, a family of sequential strategy pairs is constructed with the property that, for any positive epsilon there is a strategy pair in the family such that, for sufficiently large T, the strategy pair is a perfect (noncooperative) epsilon-equilibrium of the T-period game, and yields each player an average expected utility per period that is within epsilon of his expected utility in the one-period cooperative arrangement.

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6801

411 099

Unclassified
Security Classification

3ND PPSO 13152